

## ABSTRACT OF THE DISCLOSURE

One embodiment of the present invention provides a method and system for  
5 selecting a subset of normalization features, or biomolecule probes, from multiple data sets  
generated from a single microarray and from multiple data sets generated from multiple  
microarrays. In this embodiment, the signal intensities corresponding to common features  
within the data sets are viewed as generating a distribution of features within an  $n$ -  
dimensional signal-intensity distribution. One or more order-preserving sequences of features  
10 within the  $n$ -dimensional signal-intensity distribution are determined using an efficient, two-  
pass-per-dimension method. Normalizing features then selected from the one or more order-  
preserving sequences. Normalizing data points from generalized data sets may be obtained  
by using embodiments of the present invention.